Birnholtz, J. P., and Horn, D. B. (2007). Shake, rattle and roles: Lessons from experimental earthquake engineering for incorporating remote users in large-scale e-science experiments. *Journal of Computer-Mediated Communication*, *12*(2), article 17. http://jcmc.indiana.edu/vol12/issue2/birnholtz.html

Shake, Rattle and Roles: Lessons from Experimental Earthquake Engineering for Incorporating Remote Users in Large-Scale E-Science Experiments

Jeremy P. Birnholtz
University of Toronto
Knowledge Media Design Institute

Daniel B. Horn
U.S. Army Research Institute

Go to a section in the article:

Abstract

While there has been substantial interest in using e-science and cyberinfrastructure technologies to enable synchronous remote participation in experimental research, the details of such participation are in question. On the one hand, there is a desire to give remote participants the same views and capabilities that they would have as local participants. On the other hand, there are settings where experimental specimens and apparatus are large and difficult to manipulate effectively or view from a remote vantage point. This article argues for more novel forms of remote participation by drawing on exploratory interview and observation data gathered in civil engineering laboratories. It is shown that, while experiments are in progress, the engineers studied focus primarily on detecting and preventing specimen failures, and that their unease about remote participation stems from doubts about the ability of remote participants to detect failures adequately. It is argued that this presents the opportunity to consider novel roles for remote participants that exploit the features of e-science technologies.

Report Documentation Page

Form Approved
OMB No. 0704-0188

1. REPORT DATE **2007**	2. REPORT TYPE	3. DATES COVERED **00-00-2007 to 00-00-2007**

4. TITLE AND SUBTITLE **Shake, Rattle and Roles: Lessons from Experimental Earthquake Engineering for Incorporating Remote Users in Large-Scale E-Science Experiments**	5a. CONTRACT NUMBER 5b. GRANT NUMBER 5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S)	5d. PROJECT NUMBER 5e. TASK NUMBER 5f. WORK UNIT NUMBER
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) **U.S. Army Research Institute ,2511 Jeff-Davis Hwy,Arlington,VA,22202-3926**	8. PERFORMING ORGANIZATION REPORT NUMBER
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)	10. SPONSOR/MONITOR'S ACRONYM(S) 11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release; distribution unlimited

13. SUPPLEMENTARY NOTES

14. ABSTRACT
While there has been substantial interest in using e-science and cyberinfrastructure technologies to enable synchronous remote participation in experimental research, the details of such participation are in question. On the one hand, there is a desire to give remote participants the same views and capabilities that they would have as local participants. On the other hand, there are settings where experimental specimens and apparatus are large and difficult to manipulate effectively or view from a remote vantage point. This article argues for more novel forms of remote participation by drawing on exploratory interview and observation data gathered in civil engineering laboratories. It is shown that, while experiments are in progress, the engineers studied focus primarily on detecting and preventing specimen failures, and that their unease about remote participation stems from doubts about the ability of remote participants to detect failures adequately. It is argued that this presents the opportunity to consider novel roles for remote participants that exploit the features of e-science technologies.

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT **unclassified**	b. ABSTRACT **unclassified**	c. THIS PAGE **unclassified**	**Same as Report (SAR)**	**33**	

Standard Form 298 (Rev. 8-98)
Prescribed by ANSI Std Z39-18

Introduction

The facilitation of increased access to scarce research apparatus and resources was among the first of many potential benefits of e-science (Nentwich, 2003; RCUK, 2006) and cyberinfrastructure (Atkins, et al., 2003) technologies to be explored (Finholt, 2003; NRC, 1993). Consequently, a range of collaborative projects have sought to increase access to and aggregate data from remote shared instruments (e.g., Olson, et al., 1998) and to provide limited remote manipulation capabilities for small-scale experimental apparatus, such as microscopes and other lab instruments (Kouzes & Wulf, 1996; Potter, et al., 2001; Sonnenwald, Kupstas-Soo, & Superfine, 1999). What few have explored, however, are research scenarios in which the experimental specimens and apparatus are too large to be controlled effectively or observed solely by remote participants. Such settings provide both the challenge of facilitating remote access to these apparatus and the opportunity to enable new forms of participation in e-science research.

In this article, we report on our involvement in a cyberinfrastructure project that aimed to interconnect large-scale structural earthquake engineering (EE) laboratories. As will be detailed below, researchers in these labs work with experimental specimens that can be as large as five-story buildings and can take months or even years to construct. The hydraulic actuators and other apparatus used to test the specimens are expensive, rare, and exert forces that are, by definition, destructive and can be extremely dangerous if misused (Sims, 1999). This research area and others like it present an interesting puzzle for e-science. On the one hand, the scarcity of laboratory facilities strongly suggests the value of using network technologies to increase access by researchers at "peripheral" universities to laboratories at a small number of "core" universities. On the other hand, though, the scale and potential danger of the research seem

anecdotally to lead many engineers to reject outright the idea of serious researchers participating remotely in laboratory research.

One possible solution to this problem is to reconsider the nature of remote participation. In other words, if we accept as a given that participating remotely does not offer the same level of fidelity as being physically present, are there ways to leverage technologies to provide novel and perhaps privileged capabilities to remote participants? To explore this possibility, we conducted interviews and observations to better understand the nature of EE research. Our results suggest that using subtly perceived local cues to predict failures is a critical component of conducting earthquake engineering experiments, and participants' concerns about remote participation were closely linked to their perceived ability to predict potential failures from a distance. The data presented here suggest some novel and potentially important means for enabling participation, which are discussed in the final section of the article.

Background

One of the overarching goals of e-science and cyberinfrastructure programs is to enable new forms of geographically distributed collaboration (Atkins, et al., 2003; Hey & Trefethen, 2005; Nentwich, 2003). Such distributed collaborations can take many forms, ranging from asynchronous collaboration with shared computational and database resources (Buetow, 2005; Foster & Grossman, 2003; Foster & Kesselman, 1999) to synchronous remote participation or, as Benford, et al. (1998) refer to it, "teleparticipation" in physical lab experiments. Within this latter category, degree of participation can vary from passive observation to active manipulation of physical specimens (e.g., Potter, et al., 2001;

Sonnenwald, et al., 1999). How to enable effective remote participation, however, remains an open question.

Two general approaches to have been adopted in designing remote participation technologies. Some seek to provide as close an approximation as possible of " being there," while others seek to leverage remote participation technologies to create novel opportunities and experiences that would not otherwise be possible.

Approximating Being There

The first and most common approach to remote participation seeks to approximate, as closely as possible within technology and cost constraints, the experience of actually "being there." for remote participants. In the simplest case, a single networked video camera can provide views to passive observers (Postek, Bennett, & Zaluzec, 1999), and some basic camera manipulation can be provided.

Another simple way to move beyond passive observation is to use a combination of video or other data views and lightweight chat (e.g., Birnholtz, Finholt, Horn, & Bae, 2005; Olson, et al., 1998). This allows for remote observers to ask clarification questions or provide suggestions in a relatively unobtrusive way, although someone at the physical site must monitor the chat stream. Voice communication in combination with video can circumvent this problem, although this works less well in large groups due to coordination issues such as turn-taking and floor control (Kraut, 2003). Others have experimented with the provision of physical robotic avatars that can be controlled by a remote participant and that include cameras and other communication functionality (e.g., Jouppi, 2002; Paulos & Canny, 1998).

Where more active participation is desired, in the form of direct manipulation of equipment, specimens and parameters,

various interfaces allow for simultaneous observation and control of instruments. The Nanomanipulator, for example, allows remote participants to conduct detailed observation and manipulation of nano-scale materials and objects (Superfine, Falvo, Taylor, & Washburn, 2002). Other examples that allow for manipulation of specimens include the Bugscope (Potter, et al., 2001) and Chickscope (Bruce, et al., 1997) projects that provide schoolchildren with access to research equipment.

One common trait shared by many of these systems is their focus on small objects that can be seen in a single screen, or resources, such as the nanoscale objects in the Nanomanipulator, that are so small that they would need to be viewed on a screen even locally. Where these conditions do not hold, there is evidence to suggest that enabling remote participation may be more complicated. Gaver, et al. (1993), for example, noted that the availability of multiple views in a media space can distort how the views relate to each other and what the physical configuration of the space is actually like. Fussell, et al. (2003) further demonstrated that when multiple views are available, participants rarely take the time to select among them, even when there are clear benefits to doing so. Moreover, pan-tilt-zoom functionality increases field of view but can be distracting, and participants seldom bother to use it (Ranjan, Birnholtz, & Balakrishnan, 2006).

This leaves the question of how to facilitate effective remote participation in large-scale laboratory research. In the next section of the article, we describe a second approach to remote participation that has not yet been explored adequately in e-science. Based on data described below, we argue that such an approach holds significant promise for large-scale laboratory research.

Beyond Being There

In a now classic paper in the Human Computer Interaction literature, Hollan and Stornetta (1992) argued that seeking to approximate "being there" is a potentially debilitating constraint on the design process for remote participation technologies. Even the best video and audio links offer constrained views and are limited to what can be captured effectively by cameras and microphones. A range of studies have suggested that such attempts at "presence" are not optimal for the completion of many types of tasks (Olson & Olson, 2001). Hollan and Stornetta's critical point, is that designing technologies to provide only a "being there" experience reduces the probability that technologies will be used to facilitate novel forms of remote participation that would not be possible with only collocated participants. In other words, designers need to think beyond replication and toward innovations that exploit the unique attributes of the technologies being used.

In bringing this general approach to bear on remote participation technologies for e-science, there have been some noteworthy examples of asynchronous participation have emerged. NASA's ClickWorkers program, for example, made use of thousands of amateur space-enthusiast volunteers in effectively identifying craters in a massive set of Mars photographs (Kanefsky, Barlow, & Gulick, 2001). In the commercial realm, Amazon's Mechanical Turk (Amazon, 2006) site goes beyond volunteers and pays people to do Human Intelligence Tasks, those that are difficult for computers but relatively easy for people. Similarly, Von Ahn and colleagues (2004) have developed games played by distributed groups of participants to aid in tasks such as labeling online images. All of this suggests that there is potential value in enabling novel forms of distributed participation in e-science, although it leaves open our initial question of how to accomplish this

for synchronous participants in large-scale laboratory experiments. There have been few examples of effective remote participation in such work.

This leads us with to an interesting puzzle. Given that our goal to move beyond an approximation of "being there" for earthquake engineering experiments, we wondered what goals the engineering researchers had that could be facilitated by remote participation. This turns out to be a nontrivial question. Analysis and discovery typically take place after data are gathered and cleaned, so this does not seem like an appropriate goal for synchronous remote participants during the experiment itself. What, then, are engineering researchers doing during their experiments? Why would they want to participate remotely or have remote participants involved? These are the questions we seek to answer in this exploratory study.

Research Setting and Methods

Experimental Earthquake Engineering

Structural earthquake engineering (EE) research is concerned with understanding the responses of materials and structures to seismic forces. Work consists of field evaluation of structures, numerical simulation, and laboratory tests. Our study here is primarily concerned with how laboratory tests are conducted.

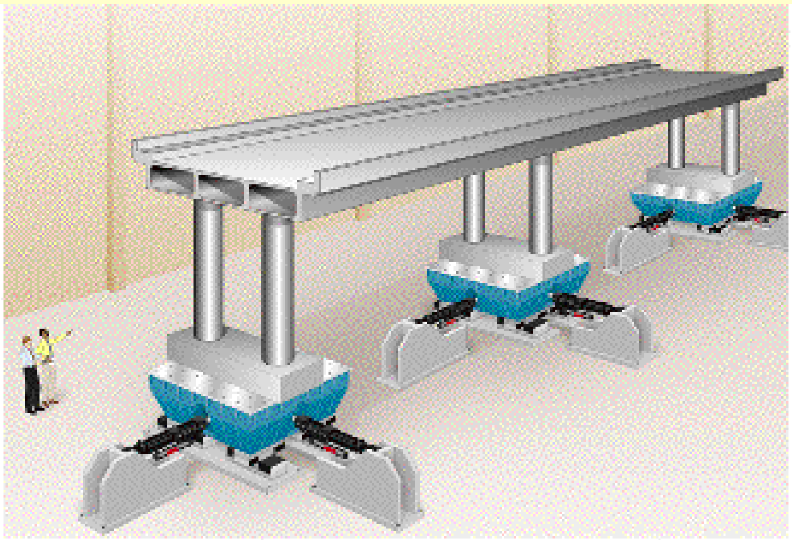

Figure 1. Artist's rendition of a full-size bridge deck that spans the three shaking tables in the structures lab at the University of Nevada, Reno (image courtesy of the University of Nevada, Reno)

NEES: Cyberinfrastructure for Earthquake Engineering

In a typical lab test, a full-size or scale model of a real-world structure is constructed, instrumented with sensors, and placed on a large testing apparatus, such as a concrete "strong wall" or large shaking platform (Sims, 1999). Graduate students take several weeks or months to build the specimen under the supervision of faculty and technicians. As Figure 1 illustrates, specimens are typically constructed of steel or reinforced concrete and range in height from five or six feet to full-scale, multi-story buildings. The specimen is then subjected to a series of pre-orchestrated, increasing stresses, which reproduce ground motion from actual earthquakes at various scales, until the specimen experiences structural failure. Given the scale of these experiments and the use of materials like concrete and steel, unexpected failure of the testing equipment or the specimen itself can be dangerous and waste large amounts of money and effort.

We studied this community in part because our research team was invited to specify the user requirements for The George E. Brown, Jr. Network for Earthquake Engineering Simulation (NEES) (NEES, 2006), a National Science Foundation project aimed at improving research, education, and practice in EE. In addition to developing a computing

infrastructure to enable collaboration among researchers, educators, and practitioners, the NEES project includes funding for constructing or upgrading EE testing equipment at 15 universities across the United States.

This computing infrastructure, called NEESgrid, includes a data repository, collaboration tools, and some support for remote participation in earthquake engineering tests. System design and programming were carried out by a multi-university group advised by our research team. The design process for NEESgrid roughly followed a spiral model (for a more detailed description of the design and management of the project, see Finholt & Birnholtz, 2004). Although there were many design decisions to be made for each component of this system, this article focuses on remote participation in experiments.

Methods

Between October 2000 and October 2003, members of our team visited 15 universities that received NEES-funded equipment and one that did not. Although we visited primarily NEES-funded sites, our sample is representative because it includes most of the 22 U.S. universities that support a student chapter of the Earthquake Engineering Research Institute, the primary professional organization in this field. Each visit consisted of a tour of the laboratory and interviews with faculty, students, and staff. We also observed tests in progress at three of the sites.

Interviews

We interviewed 94 subjects at 14 sites (see Table 1). Subjects included faculty, students, and technicians. All interviews were tape-recorded and typically conducted by two members of our team: one person asked questions while the other

took notes. The note taker typed full transcripts after each interview, consulting the audiotape when details were unclear.

As our larger goal was the specification of user requirements for NEESgrid, our questions focused on the process of conducting investigations, from idea to published paper. As we conducted more interviews, the protocol iteratively became more structured and focused. The same protocol was used for all subjects, but emphasis on specific issues was shifted based on the subject′s experience and expertise. Where individuals are directly quoted below, each is identified with a unique letter, which corresponds to his or her title listed in the Appendix.

	Number Interviewed			Total at Site		
Site	Faculty	Students	Technicians	Faculty	Students	Technicians
1	6	0	2	15	10	2
2	6	1	2	18	10	1
3	4	2	2	9	10	3
4	4	2	0	6	6	0
5	4	3	1	14	27	2
6	3	2	2	6	3	2
7	3	2	1	12	3	1
8	4	0	1	13	12	2
9	1	4	1	9	14	4
10	6	9	2	9	16	3
11	1	3	2	7	6	2
12	4	2	1	9	15	2
Total	51	28	15	134	143	27

Table 1. Number and type of interviews conducted and total number of earthquake engineering researchers present at all sites.

Observation

At three of the sites, we observed tests in progress (Lofland & Lofland, 1995). These observations were helpful in furthering our understanding of work processes in the labs we studied. Observations consisted of being in the laboratory setting for the duration of the setup and execution of the experiment, which ranged from a few hours in one case to three full days in another. Field notes were taken during meals, breaks, and at the end of each day. When possible, the observers assisted with routine laboratory tasks and asked the researchers to explain what they were doing.

Analysis

To analyze the data, we used inductive qualitative techniques (Huberman & Miles, 1994; Miles & Huberman, 1994), while keeping our focus on remote participation and failure prediction. Data analysis consisted of reading and rereading interview notes and field notes and examining photographs, videotapes and other artifacts. This focus guided another examination of our data in which we further developed the themes and assessed the extent to which they accurately represent laboratory practices, based on both our subjects' descriptions and our own ethnographic observations. We also discussed our preliminary findings with a small number of earthquake engineers to generate additional feedback, discussion, and insights.

Findings: Current Practices and Beliefs

In our analysis, two themes emerged as particularly prevalent: *failure prevention* and *concerns about remote manipulation*.

Failure Prediction During Experiments

Specimen failure is most likely to occur very early or very late in the testing process. Interestingly, failures that occur early in the testing process are always undesirable, while only some failures late in the testing process are undesirable. This is because early failure is typically a sign of a flaw in the design or implementation of the specimen or testing apparatus and occurs before the desired data have been collected. Late failures, on the other hand, occur after such data have been collected, and the data collected during failure is often part of the planned testing protocol. As a specimen

nears its predicted point of failure, however, it could succumb to the forces exerted by the equipment earlier than expected or in an unpredicted fashion. Thus, there is a strong desire both to exert sufficient force on the specimen so that it fails (collapses) and retain sufficient control so that it fails in a safe manner. As one of our interview subjects noted, " the goal is destruction of the specimen, not the testing apparatus.[B]"

When asked to describe what they do during a test, all subjects but one mentioned that they look for signs of potential failures. Respondents also reported that not knowing where failures might come from mandates vigilance and, many believed, physical presence in the laboratory. For example, one subject said, "when I have a specimen being tested, I am extremely cautious and nervous. I just cannot imagine not being there for the test.[A]" Another said, "I think people need to be here. I need to see what's going to happen. [B]" In exploring the details of how local failure prediction occurs, we noticed three themes that are elaborated below: 1) the use of multiple sensory cues, 2) variable likelihood of failure, and 3) integration of multiple viewpoints.

Sensory cues involved

One theme that we observed in exploring our data on failure prediction is that earthquake engineers tend to integrate multiple sensory and information streams in the process of predicting possible failures during experiments. Subjects indicated that they regularly relied on multiple information sources during an experiment. The three sources most commonly mentioned were (in order of frequency of mention): 1) onscreen displays of numerical data, 2) walking around and looking at the specimen, and 3) sounds being made by the testing apparatus and the specimen. One subject also reported that he goes by the "feel" of the vibrations given off by the shaking table in his laboratory.

Most of the subjects reported looking at numerical or graphical displays of data from sensors and instruments on the specimen itself, and we confirmed this to be true in our observations. Subjects looked at these data displays to ensure that all the sensors were working properly. As one subject indicated, "I want to make sure the instruments are working, that the data are coming in and being recorded.[C] " In light of the costs in terms of both time and money associated with experiments, the importance accorded to data integrity is not surprising. Subjects also reported looking at the data to make sure the experiment was progressing as expected and that there were no extreme anomalies. This is typically accomplished by looking at a chart of force (or stress being placed on the specimen) versus. displacement (the degree to which the specimen is moving). One subject noted that on his tests, "if we can't explain the graphs, we stop immediately. If we get data that are surprising, but not crazy we'll keep going.[E] " The interesting implication here is that experiments necessarily involve some uncertainty, but there appears to be a significant and deliberate effort to mitigate risk by detecting anomalous behavior and determining whether it is within the scope of the investigation and potentially informative ("surprising") or evidence of a potential failure that might be present in the system that must be detected ("crazy").

Most subjects also reported looking at the specimen to predict failures and spot potential trouble, and we also found this to be true in our observations. One subject said that, "we are examining the specimen itself, looking around for visible signs of distress, like cracking.[F]" This is frequently combined with looking at the numerical data to supplement understanding of what is taking place. One subject provided a valuable description of moving between these information

sources:

> I look at the force vs. displacement plot, because a change in slope on this plot means that something significant is going on. Next, you have to figure out where, how, and why this is happening. You do this by walking around and looking. ... [When there are problems], you can see it right away, it's very hard to hide.[G]

Thus, the integration of numerical data and visual inspection of the specimen can supplement each other.

We also found that some subjects reported relying on hearing the test in order to predict failures. Hearing was typically integrated with viewing onscreen data and looking at the specimen. One subject, a laboratory manager at a large facility, indicated that, "I go by sound when there is a problem, and then look at the specimen visually and at the numeric data.[H] " Another subject, who was also involved in field testing, noted that:

> "In a typical field test, for example, of buildings, you'll be able to hear cracking. If that's an owner-occupied building, that's real bad. That would probably put an immediate end to the test."[I]

The point here is that hearing sounds clearly plays an important role in failure prediction.

There is some evidence to suggest that subjects with more experience in EE testing are better able to understand and integrate multiple sources of information, particularly auditory information. The only people who mentioned the use of auditory information were those who had prior experience, and those rely on it as a primary or very important source of information were those with several years of experience. Additionally, one student's comment suggests that there is some difficulty in distinguishing signal from noise in an auditory information stream:

> Even after [we had fixed a problem with the test setup], there was still a lot of noise. I might have pushed the [emergency stop] button. It was very noisy.[J]

Although hearing appears to be the most difficult sense to use in failure prediction, those who have experience using hearing to predict failure appear to rely on it more heavily than they do other senses. In light of evidence from the psychological literature on expertise (e.g., Chi, Glaser, & Farr, 1988), this is not surprising. Expert workers in a range of settings have been observed to rely on subtle nuances and cues that would be difficult for novices to understand (Mumaw, Roth, Vicente, & Burns, 2000; Zuboff, 1988).

Variable likelihood of failure

Because of their experience, we would expect faculty members and technicians to be the best-equipped individuals in a lab to detect potential failures. In some labs, only technicians are permitted to control the testing equipment, so they are always present during experiments. Faculty members have more demands on their time but indicated the importance of their presence at tests to help predict failures. Because they frequently cannot be present for the entire test, we would expect them to be present when it was most likely that a potential failure would be spotted. We therefore asked faculty if they typically attended entire tests and asked their students during what parts of the tests faculty were present. Responses indicated that faculty typically showed up only for the first few and last few shaking events on a specimen. This is closely related to the belief that, as we mentioned above, failures tend to occur early and late in the tests. One subject indicated, for example, that:

> I'm always there for the first test on a particular specimen, because I need to train the students on the things they need to do ... like making sure the test frame is not creating a physical anomaly. Students have a tendency to just roll forward without checking these things.[C]

Similarly, many faculty members indicated that they are not present for the bulk of the tests on a specimen. One subject

said that she is, "not physically there watching the whole time, certainly not.[D]" Another said that, "after a while I gain confidence. I'll just show up to see what's going on and then leave.[C]"

Multiple collocated persons

The third and final theme we observed in our failure prediction data is a reliance on multiple co-located persons, both in detecting failures and in making decisions about how to prevent them. The presence of multiple persons at any test has its origins in safety concerns. Virtually all labs have a strictly enforced safety policy stating that no testing equipment may be used when fewer than two people are present. This has the effect that multiple people are involved in making the crucial decisions about how the experiment is to move forward. Our data indicate that this is valuable in two ways: 1) multiple interpretations of events and 2) the ability to make collective decisions.

First, one senior faculty member pointed out that multiple people in the lab means that "there are different accounts of what happened, like people's reports at the scene of a car accident.[K]" Integrating these multiple human sources of information can increase the clarity and understanding of what is taking place in the test.

Second, we found countless examples of informal meetings, what one subject referred to as "powwows," in the lab, in which the students, technicians and faculty members decided together how to proceed:

> When things go awry, we tend to powwow in the lab. There are usually multiple professors, we meet in the control room with [the lab manager] and the student, and try to sort out what's going on.[B]

This is valuable in that it allows for the integration not only of multiple perspectives on unfolding events, but also multiple forms of expertise. Multiple forms of expertise also enable some specialization during the experiment. One senior

technician reported that he would "often send somebody out to stand in a particular place and keep an eye on things.[L]" Another subject, a student, suggested that he likes to have "one other person around to mark cracks, take pictures, [and] take notes.[M]" Many subjects we spoke with also indicated that they participate in the "powwow" and have a significant amount of influence on what takes place, but often defer final authority to the laboratory technician, who is typically the most experienced with the test equipment. It is through collective awareness and sensitivity, combined with communication among co-located parties, that potential failures can be detected and prevented during tests.

Concerns about Remote Manipulation

A very small number of our subjects had prior experience with remote participation technologies, so this section relies considerably on their beliefs and concerns. When asked about the potential role of remote participation during tests, our respondents expressed a significant amount of interest in remote observation but not remote manipulation of the experimental apparatus. Of the subjects with whom we discussed remote participation, only one indicated that he hoped people could remotely manipulate the apparatus. Concerns about remote manipulation focused heavily on safety issues. One respondent, for example, said:

> I've heard that they actually want you to be able to run your experiment from far away. That doesn't seem possible to me … I can't imagine that [the lab manager] would be okay with someone else running the controls with him not being here. There are safety issues.[D]

Another suggested that safety is a significant issue, but remote manipulation "could be offered for some specific, limited things.[Q]" Additional concerns about remote manipulation reflected the belief that physical presence at a test provided

richer access to the "details of the test" that could not be replicated.

There was a great deal of interest, however, in remote observation of tests. Several subjects, for example, suggested that they would "tune in" to watch their colleagues' tests in progress while they worked on other things in their office. Another suggested that remote observation would make more of the shaking events comprising a test more accessible to busy faculty members because "it's hard to pin down [in advance] the exact time a shake will occur.[R] " The implication is that, with remote participation, the faculty member could watch a test in progress without having to visit the lab (which is often located in another building or even off-campus). The general consensus was that remote observation would make tests more accessible to students, sponsors, and researchers interested in secondary data analysis. In these instances remote observation provides benefits to individuals who would otherwise not have access to a test, but aside from increased publicity, there is no real advantage for those conducting the tests. A few comments indicated that remote observation could provide benefits to the research team. One researcher indicated that, "you will be able to broadcast a test and get feedback from others.[O]"

To summarize, earthquake engineers see remote participation as a way to make experiments more accessible to a wider audience, yet they are wary of giving control to those who are not physically attending a test. In our interviews, participants consistently conceptualized remote participation systems in terms of existing practices. There was little mention of new roles for remote users beyond passive observation. In the next section, however, we present an interesting case in which a researcher with many of these same expectations had an unexpected experience in which a remote observer became a valuable part of his experiment.

Empowering Remote Participants: A Pleasant Surprise

Only one subject we spoke with had previous experience with active remote participants in one of his tests. Although his story admittedly does not reflect standard practice in EE, we believe it demonstrates one novel means for remote participation that could aid in failure prediction. This researcher was a junior faculty member conducting an experiment that was part of a large, multi-institution project. During one test, he decided to implement a simple remote participation system that did not interfere with his existing work practice:

> I did it in a very crude way I just did a web cam where I uploaded images, I think every 40 seconds. Because there were people who had very slow Internet connections, so that was enough for them . . . I took all my notes in html because you've got to write down anyhow, and have a place where you can type that in is good so people can see that automatically at the same time. And I just had my email box opened up.[N]
>
> This setup was designed to enable relatively crude remote observation, but this crudeness turned out to have an unexpected benefit when a remote colleague noticed a potential anomaly and sent an email to the researcher in the lab:
>
> she noticed a lot of yielding in one area that I had the camera kind of overall and I didn't even notice it. So if she didn't tell me, I probably would have missed some of the issues that were going on.[N]

This anomaly, it turns out, was apparent to the remote observer precisely because of the crudeness of the setup. The specimen "yielding" was very slight and difficult for local participants to perceive. When the webcam video frames changed, however, the long interval between frames made the yielding much easier to spot. Moreover, the researcher reported an additional way in which this setup was useful. His use of mail as a means for interaction during this test meant that "you don't have 20 people yelling at you at the same time...and you can just look at the emails that are really important and work that way.[N]"

Discussion and Implications

We began with the goal of considering alternative forms of synchronous remote participation in EE laboratory experiments. Our data suggest strongly that EE researchers' concerns about remote participation in experiments stem from what they perceive as an inability for remote participants to use their existing expertise and methods to spot early signs of premature specimen failure. This raises two questions with important implications for our goal. First, are our data best explained by the scale and nature of EE work, and do they consequently apply only to large-scale laboratory experiments, or are they more generally representative of experimental research? Second, how might we use these findings to reconceptualize remote participation?

Explaining the Findings

As for the first question, our subjects were concerned about remote participation because of concerns about the physical safety of researchers, and they were skeptical about the capacity of teleparticipation technologies to provide adequate information for remote researchers to detect and predict small signs of specimen failure.

Safety

On the one hand, concern for participant safety is not unique. Most laboratories have safety rules and procedures, thanks in part to catastrophic failures of some early instruments (Galison, 1997). This history of Particle Physics details some of the early laboratory disasters that led to calls for improved safety. Laboratory danger, however, takes many forms. Threats to safety in many labs, for example, come from potential exposure to harmful substances (e.g., laser beams, caustic chemicals, radiation or noxious fumes). In these cases, there is value in remote manipulation technologies even

for local researchers , in that the probability of harmful exposure can be reduced.

On the other hand, EE research is unique in that our subjects were concerned not about exposure to harmful substances, but about experimental specimens and destructive forces large and powerful enough to injure or kill people and cause irreparable damage to laboratory equipment. Thus, the relationship between teleparticipation and safety is complex and deserving of future exploration. In large-scale research such as that observed here, researchers are concerned about harm that may come to local observers and equipment as a result of poor decisions made by a remote operator. This is potentially quite distinct from other research where exposure to harmful substances can actually be avoided via remote participation. To be sure, these differences in safety concerns are not exclusively a function of research scale. Based on the data presented here, however, scale is one dimension of potential consideration in assessing teleparticipation possibilities in experimental research.

Detecting and Predicting Specimen Failure

Our subjects were also concerned about being able to detect signs of potential specimen failure when participating remotely. Based on their descriptions of current practice, the ability to detect subtle, local cues and rely on multiple interpretations of events emerged as particularly important in our data. Concern for experimental integrity and the use of subtle cues in assessing threats to such integrity are not unique. All researchers want their experiments to succeed, and others have demonstrated that experience with physical apparatus, experimental or otherwise, leads to the aggregation of tacit knowledge which is then utilized in procedures that may be difficult to articulate and implement in

teleparticipation scenarios (Collins, 1985; Zuboff, 1988).

At the same time, however, there are key points of distinction between the researchers we observed and others. First, the cost of experiment failure varies widely across fields. Computer scientists, for example, frequently use a spiral model of software development (Boehm, 1995) because failure is expected, and the code is easy to edit or repair when it does fail. At the other extreme, the specimens in the laboratories we visited can cost hundreds of thousands of dollars, take months or years to build, and cannot be repaired in the event of premature failure. Thus, the high cost of failure appears to be one motivator in the wariness of our subjects about remote manipulation of experimental apparatus.

This leads to the second point of distinction, which is more closely linked to scale. In the laboratories we visited, specimens were large and difficult to capture completely in a small number of video views. Standing in the laboratory with the specimen affords a more complete view and the ability to move quickly and intuitively to any location for more narrowly targeted observations. This is notably distinct from work in other areas where, for various reasons, the specimen is not directly observable (e.g., electron microscopes, particle accelerators, centrifuges). In these cases, there is no advantage to being in the laboratory with the specimen because the experiment must be observed using cameras anyway. These are, admittedly, extremes, but such cases does suggest that the direct observability of the specimen is an important factor in considering people’s ability to predict failure.

Thinking about the Future

These findings of this study suggest several guidelines in considering how to reconceptualize remote participation in experimental research. While it may be natural for people to think of remote participation facilities in terms of providing a

low fidelity imitation of the environment that individuals experience when they are physically present, there are cases where benefits may be garnered from a different representation of the problem. In certain contexts, a "beyond being there" approach, in which remote observation tools are designed to *complement* the information that is available to those who are attending a test, could be quite useful. To make such an approach work, teleparticipants would have to be able use information that physically present observers cannot or do not use. This could result in improved research quality and greater enthusiasm for remote participation technologies.

One example would be to implement filters that highlight features of interest on streaming\ video. Physically present observers are not likely to rely on streaming video given that they can directly observe the specimen in front of them. For example, if teleparticipants could view video of a live test with overlays indicating visual features that are difficult to discern in person, such as out-of-reach portions of a specimen, or equipment that moves slowly, they would be able to offer different observations than additional physically present observers could. It would, of course, be possible to provide similar video views and filters to a co-present observer, but physically present observers will already be occupied by a great deal of higher-fidelity sensory information, making it difficult to attend to additional views. Even where local participants could be assigned to watch a particular video view, our data suggest that there are few "spare" local participants, and spatial constraints prohibit large numbers of observers. There are no such constraints for remote participants, which raises the problem, of course, of integrating input from large numbers of remote observers in ways that extract valuable information without distracting local participants. These types of problems are being addressed by

other researchers (Amazon, 2006; Bisantz, Finger, Seong, & Llinas, 1999; Kanefsky, et al., 2001; von Ahn & Dabbish, 2004).

Finally, we should note that we do not believe that failure prediction is the only reason people are wary of remote participation. Although this is undoubtedly a central aspect of EE testing practices, the success of remote participation will depend on other factors as well, such as peer acceptance and political and funding pressure.

Future Work

This work opens a range of areas for additional investigation. First, novel remote participation interfaces should be developed and implemented. While we have learned a great deal about research in EE and have made a theoretical case for novel forms of participation, these ideas must be tested and evaluated via field deployment. This means thinking carefully about how remote participants could be engaged and what views and interfaces would be useful to the primary researchers. We have provided some preliminary suggestions above, but these are open to further development.

Second, there is value in thinking carefully about the goals of e-science collaborators at all phases of the research process. Failure prediction has little to do with earthquake engineering as a research discipline, yet it appears to be an extremely important skill in conducting effective laboratory research. In designing technologies and allocating resources for e-science, it is important to think not just about the long-term, high-level goal of enabling novel discoveries, but also about how technologies can be leveraged to support short-term goals during different phases of the research process. More research is needed to determine the extent to which these findings generalize to other research fields and to investigate the nature of researchers’ goals at other phases in the research process.

Acknowledgments

This work was supported in part by the National Science Foundation (CMS 0117853). We wish to thank the members of the Earthquake Engineering community for their participation. In addition, we thank Mark Ackerman, Matt Bietz, Jane Dutton, Tom Finholt, Christine Halverson, Nick Jankowski, Karl Weick, and several anonymous reviewers for their insightful comments on earlier drafts of this work.

Notes

1. The author listing is alphabetical. This was an entirely collaborative effort.
2. The views, opinions, and/or findings contained in this article are solely those of the authors and should not be construed as an official Department of the Army or DOD position, policy, or decision, unless so designated by other documentation.

References

Amazon. (2006). *Mechanical Turk: Artificial Artificial Intelligence*. Retrieved May 15, 2006 from http://www.mturk.com/mturk/welcome

Atkins, D. E., Droegemeier, K. K., Feldman, S. I., Garcia-Molina, H., Klein, M. L., & Messina, P. (2003). *Revolutionizing Science and Engineering Through Cyberinfrastructure: Report of the National Science Foundation Blue-Ribbon Advisory Panel on Cyberinfrastructure*. Washington, D.C.: National Science Foundation.

Benford, S., Greenhalgh, C., Reynard, G., Brown, C., & Koleva, B. (1998). Understanding and constructing shared spaces with mixed-reality boundaries. *ACM Transactions on Computer-Human Interaction, 5* (3), 185-223.

Birnholtz, J. P., Finholt, T. A., Horn, D. B., & Bae, S. J. (2005). Grounding needs: Achieving common ground via lightweight chat in large, distributed ad-hoc groups. In *Proceedings of the ACM Conference on Human Factors in Computing Systems* (pp. 287-296). New York: ACM Press.

Bisantz, A. M., Finger, R., Seong, Y., & Llinas, J. (1999). Human performance and data fusion based decision aids. In *Proceedings of the 2nd International Conference on Information Fusion*. vol. 2 (pp. 918-925). Sunnyvale, CA.

Boehm, B. W. (1995). A spiral model of software development and enhancement. In R. M. Baecker, J. Grudin, W. A. S. Buxton & S. Greenberg (Eds.), *Human Computer Interaction: Toward the Year 2000* (pp. 281-292). San Francisco, CA: Morgan Kaufman.

Bruce, B. C., Carragher, B., Damon, B. M., Dawson, M. J., Eurell, J. A., Gregory, C. D., et al. (1997). Chickscope: An interactive MRI classroom curriculum innovation for K-12. *Computers Education, 29* (2/3), 73-87.

Buetow, K. H. (2005). Cyberinfrastructure: Empowering a "third way" in biomedical research. *Science, 308* (5723), 821-824.

Chi, M. T. H., Glaser, R., & Farr, M. J. (1988). *The Nature of Expertise*. Hillsdale, NJ: L. Erlbaum Associates.

Collins, H. (1985). *Changing Order*. London: Sage Publications.

Finholt, T. A. (2003). Collaboratories as a new form of scientific organization. *Economics of Innovation and New Technologies, 12* (1), 5-25.

Finholt, T. A., & Birnholtz, J. P. (2004). If we build it, will they come? The cultural challenges of cyberinfrastructure development. In W. Bainbridge & M. Roco (Eds.), *Managing Nano-bio-info-cogno Innovations* (pp. 89-102). Berlin: Springer.

Foster, I., & Grossman, R. L. (2003). Data integration in a bandwidth-rich world. *Communications of the ACM, 46* (11), 51-57.

Foster, I., & Kesselman, C. (1999). *The Grid: Blueprint for a New Computing Infrastructure.* New York: Morgan Kaufmann.

Fussell, S. R., Setlock, L. D., & Kraut, R. E. (2003). Effects of head-mounted and scene-oriented video systems on remote collaboration on physical tasks. In *Proceedings of the ACM Conference on Human Factors in Computing Systems* (pp. 513-520). New York: ACM Press.

Galison, P. (1997). *Image and Logic: A Material Culture of Microphysics*. Chicago, IL: University of Chicago Press.

Gaver, W., Sellen, A., Heath, C., & Luff, P. (1993). One is not enough: Multiple views in a media space. In *Proceedings of*

InterCHI (pp. 335-341). New York: ACM Press.

Hey, T., & Trefethen, A. E. (2005). Cyberinfrastructure for e-science. *Science, 308* (5723), 817-821.

Hollan, J., & Stornetta, S. (1992). Beyond being there. In *Proceedings of the ACM Conference on Human Factors in Computing Systems* (pp. 119-125). New York: ACM Press.

Huberman, A. M., & Miles, M. B. (1994). Data management and analysis methods. In Y. S. Lincoln & N. K. Denzin (Eds.), *Handbook of Qualitative Research* (pp. 428-445). Thousand Oaks, CA: Sage.

Jouppi, N. (2002). First steps towards mutually-immersive mobile telepresence. In *Proceedings of the ACM Conference on Computer Supported Cooperative Work* (pp. 354-363). New York: ACM Press.

Kanefsky, B., Barlow, N. G., & Gulick, V. C. (2001). Can distributed volunteers accomplish a massive data analysis task? In *Proceedings of the Lunar and Planetary Science XXXII*, # 1272.

Kouzes, R., Myers, & Wulf, W. (1996). Collaboratories: Doing science on the Internet. *IEEE Computer, 29* (8), 40-46.

Kraut, R. (2003). Applying social psychological theory to the problems of group work. In J. M. Carroll (Ed.), *HCI Models, Theories and Frameworks* (pp. 325-356). New York: Morgan Kaufmann.

Lofland, J., & Lofland, L. H. (1995). *Analyzing Social Settings: A Guide to Qualitative Observation and Analysis*. Detroit: Wadsworth Publishing Co.

Miles, M. B., & Huberman, A. M. (1994). *Qualitative Data Analysis: An Expanded Sourcebook*. Thousand Oaks: Sage Publications.

Mumaw, R. J., Roth, E. M., Vicente, K. J., & Burns, C. M. (2000). There is more to monitoring a nuclear power plant than meets the eye. *Human Factors, 42*, 36-55.

NEES. (2006). NEESinc website. Retrieved May 16, 2006 from http://www.nees.org/

Nentwich, M. (2003). *Cyberscience: Research in the Age of the Internet*. Vienna: Austrian Academy of Sciences.

NRC. (1993). *National Collaboratories: Applying Information Technology to Scientific Research*. Washington D.C.: National Research Council.

Olson, G. M., Atkins, D. E., Clauer, R., Finholt, T. A., Jahanian, F., Killeen, T. L., et al. (1998). The upper atmospheric research collaboratory. *Interactions, 5* (3), 48-55.

Olson, G. M., & Olson, J. S. (2001). Distance matters. *Human-Computer Interaction, 15*, 139-179.

Paulos, E., & Canny, J. (1998). PRoP: Personal roving presence. In *Proceedings of the ACM Conference on Human Factors in Computing Systems* (pp. 296-303). New York: ACM Press.

Postek, M. T., Jr., Bennett, M. H., & Zaluzec, N. (1999). Telepresence: A new paradigm for industrial and scientific

collaboration. *Proceedings of SPIE, 3677*, 599-610.

Potter, C. S., Carragher, B., Caroll, L., Conway, C., Grosser, B., Hanlon, J., et al. (2001). Bugscope: A practical approach to providing remote microscopy for science education outreach. *Microscopy and Microanalysis, 7*, 249-252.

Ranjan, A., Birnholtz, J. P., & Balakrishnan, R. (in press). An exploratory analysis of partner action and camera control in a video-mediated collaborative task. To appear in *Proceedings of the ACM Conference on Computer Supported Cooperative Work*, November 4-8, 2006.

RCUK. (2006). *About the UK E-Science Programme*. Retrieved May 16, 2006 from http://www.rcuk.ac.uk/escience/

Sims, B. (1999). Concrete practices: Testing in an earthquake engineering laboratory. *Social Studies of Science, 29* (4), 483-518.

Sonnenwald, D. H., Kupstas-Soo, E., & Superfine, R. (1999). A multi-dimensional evaluation of the nanoManipulator, a scientific collaboration system. *ACM SIGGROUP Bulletin, 20* (2), 46-50.

Superfine, R., Falvo, M., Taylor, R. M., & Washburn, S. (2002). Nanomanipulation: Buckling, transport and rolling at the nanoscale. In D. Lyshevski, J. Brenner, J. Lafrate & W. Goddard (Eds.), *CRC Handbook of Nanoscience, Engineering and Technology*. Boca Raton, FL: CRC Press LLC.

von Ahn, L., & Dabbish, L. (2004). Labeling images with a computer game. In *Proceedings of the ACM Conference on Human Factors in Computing Systems* (pp. 319-326). New York: ACM Press.

Zuboff, S. (1988). *In the Age of the Smart Machine: The Future of Work and Power*. New York: Basic Books.

Appendix: Quoted Interview Subjects

[A] Professor of Civil Engineering

[B] Associate Professor of Civil Engineering

[C] Associate Professor of Civil and Environmental Engineering

[D] Assistant Professor of Civil and Environmental Engineering

[E] Graduate Student, Structural Engineering

[F] Assistant Professor of Civil and Environmental Engineering

[G] Associate Professor of Civil Engineering

[H] Research Assistant Professor of Civil Engineering and Laboratory Manager

[I] NEES Site Manager

[J] Graduate Student, Structural Engineering

[K] Professor of Civil Engineering

[L] Senior Development Engineer and Lab Manager, Structural Engineering

[M] Graduate Student, Civil and Environmental Engineering

[N] Assistant Professor of Civil Engineering

[O] Graduate Student, Civil Engineering

[P] Professor of Structural Engineering

[Q] Assistant Professor of Civil and Environmental Engineering

[R] Facility Manager, Department of Civil and Environmental Engineering

About the Authors

Jeremy P. Birnholtz is a postdoctoral fellow in the Knowledge Media Design Institute at the University of Toronto. His research interests are in the area of improving communication and collaboration tools for workgroups.

Address: University of Toronto, Knowledge Media Design Institute, 40 St. George St., Room 7204, Toronto, Ontario M5S 2E4, Canada

Daniel B. Horn is a Research Psychologist at the U.S. Army Research Institute for the Behavioral and Social Sciences in Arlington, VA. He has conducted research on video-mediated communication, the usability of speech recognition

software, geographically distributed scientific collaboration, and the use of videogame technology for training.

Address: U.S. Army Research Institute, 2511 Jefferson Davis Hwy., Arlington, VA 22202-3926 USA.